

Response to RC1

Jiachen Liu¹, Eric Chen¹, Shannon L. Capps¹

¹Department of Civil, Architectural & Environmental Engineering, Drexel University, Philadelphia, Pennsylvania, USA

Correspondence to: Shannon L. Capps (shannon.capps@drexel.edu)

For the convenience of the reviewer, we have included the comments from Reviewer 1 in black, normal font style and our responses indented in green, normal font style. The changes made to the manuscript or supplemental information corresponding to our response are provided in green, bold style font with the context of line numbers from the original manuscript and the original, remaining text in the normal font style.

In this paper, Liu et al., have applied the hyper-dual sensitivity analysis approach to a chemical transport model (CMAQ). They find the method to be both accurate and computationally relatively efficient for calculating first and second order sensitivities. In general, the manuscript is scientifically sound and well written.

- First and foremost, we would like to thank the reviewer for giving careful attention to this work and for the favorable assessment of it.
- We have also corrected a miscellaneous inconsistent spelling of modeling from **lines 128 to 129**:

Lines 128 to 129: Rehner and Bauer (2021) applied hyperdual numbers to equation of state **modelling** and the calculation of critical points.

“DDM-3D” was developed by Yang et al., in 1997.

- Thank you for the correction. We have revised the **lines 93 to 96** describing the development of DDM-3D in the manuscript.

Lines 93 to 96: **On the other hand**, DDM formulates sensitivity equations like the direct method but separately solves the original and sensitivity equations. This **approach** improves the computational efficiency and stability **compared to** the direct method. **Yang et al. (1997) was the first to apply the DDM-3D method in a three-dimensional chemical transport model.**

The development of (5) needs to be further explained. Given it is a Taylor series expansion, about what value is the expansion?

- Thank you for highlighting the need for more background on applying the hyperdual step method. We have made the connection to the real-valued Taylor series expansion more explicit in the text and have directed the reader to the SI, where a derivation of the multiplication of a real number and hyperdual perturbation leading to Eq. (5) is now included.

The first- and second-order sensitivities are in the ϵ_1 , ϵ_2 , and ϵ_{12} terms. After further simplification, Eq. (5) becomes Eq. (6), and the numerically exact sensitivities are separated and listed in Eq. (7) and Eq. (8). We have also included a sentence describing Eq. (S7).

Lines 151-159: **Akin to the Taylor series expansion about the real value of x in the finite difference method**, the method of ascertaining sensitivities through a perturbation in hyperdual space is based on a Taylor series expansion in an orthogonal dimension of the number. Specifically, a hyperdual number with unity in a_0 and unity in one of a_1 , a_2 , and a_{12} is multiplied with the independent variable of interest. After model execution, a Taylor series expansion is applied to extract sensitivities. **For instance**, the hyperdual-step method is applied to a scalar function $f(x)$ by multiplying x by the hyperdual number $H_h = 1.0 + h_1\epsilon_1 + h_2\epsilon_2$, **which** results in:

$$\begin{aligned} f(xH_h) &= f(x) + (xh_1\epsilon_1 + xh_2\epsilon_2)f'(x) \\ &\quad + \frac{1}{2!}(xh_1\epsilon_1 + xh_2\epsilon_2)^2f''(x) \\ &\quad + \frac{1}{3!}(xh_1\epsilon_1 + xh_2\epsilon_2)^3f'''(x) + \dots \end{aligned} \quad (5)$$

where “...” represents higher order terms in the series. Eliminating all terms that are zero due to the definition of hyperdual numbers (Eq. 2) **leads to**

$$f(xH_h) = f(x) + (xh_1\epsilon_1 + xh_2\epsilon_2)f'(x) + x^2h_1h_2\epsilon_{12}f''(x) \quad (6)$$

where $f(xH_h)$ is a hyperdual number.

The properties of hyperdual numbers (Eqs. 2–4) lead to two significant results. First, all terms in the Taylor series expansion with derivatives higher than second-order become zero because all values include ϵ_1^2 , ϵ_2^2 , or ϵ_{12}^2 . Second, the real component is unchanged. **A more detailed expansion of terms can be found in Eq. S7 in the SI or the original development of hyperdual numbers, following the multiplication rule between a hyperdual and a real number** (Fike and Alonso, 2011).

Line 15, SI:

$$\begin{aligned} f(xH_h) &= f(x * (1.0 + h_1\epsilon_1 + h_2\epsilon_2)) \\ &= f(x + xh_1\epsilon_1 + xh_2\epsilon_2) \\ &= f(x) + (xh_1\epsilon_1 + xh_2\epsilon_2)f'(x) + \frac{1}{2!}(xh_1\epsilon_1 + xh_2\epsilon_2)^2f''(x) + \frac{1}{3!}(xh_1\epsilon_1 + xh_2\epsilon_2)^3f'''(x) + \dots \\ &= f(x) + (xh_1\epsilon_1 + xh_2\epsilon_2)f'(x) + \frac{1}{2}(x^2h_1^2\epsilon_1^2 + 2x^2h_1\epsilon_1h_2\epsilon_2 + x^2h_2^2\epsilon_2^2)f''(x) \\ &= f(x) + xh_1\epsilon_1f'(x) + xh_2\epsilon_2f'(x) + x^2h_1h_2\epsilon_{12}f''(x) \end{aligned}$$

The Taylor expansion of the multiplicative hyperdual perturbation (Eq. 5) is shown in Eq. S7.

In (16) and (17) is E_{Nox} a function of space or time? If yes, the derivatives calculated are very complex, and indeed, it would be good for the authors to explain exactly how they are taking those derivatives and how the set of mathematical operations are being done. If E_{Nox} is not space or time dependent, what is it (how is it mathematically defined)? I think I know what they are trying to do, but the current representation needs to be clarified and made mathematically more precise. They should indicate the spatial and temporal dependencies in the variables.

- Thank you for this question and helpful comment. In responding to it, we noticed and corrected some errant numbering for the equations in the original manuscript. There were

two occurrences of Equation (2) in the original manuscript from **lines 145 to 150**. We have revised the manuscript to correct this mistake.

- We have also corrected the numbering of Eqs. 15–19 from **lines 252 to 281**, which were mistakenly labeled as Eqs. 16–19, with two occurrences of Eq. 19. We have added to now Eqs. 15-17 the subscripts that formerly were only introduced in what is now Eq. 18.

Lines 236-250: Here, **for the sake of illustration**, we consider the semi-normalised sensitivities of **time-averaged** output concentrations on the ground layer, $l=0$, to input emissions averaged over time, t , **for any given cell as indicated by the column, c , and row, r** . First-order semi-normalized sensitivities, $S_{NO_x}^{PM_{2.5}}$, and second-order semi-normalised sensitivities, $S_{NO_x}^{(2)PM_{2.5}}$, of **ground-level** $PM_{2.5}$ concentrations, $C_{PM_{2.5},c,r,l=0,t}$, to NO_x ($NO+NO_2$) emissions, $E_{NO_x,c,r,l,t}$, exemplify sensitivities relevant to environmental decision makers (Eqs. **15–16**).

$$S_{NO_x}^{PM_{2.5}} = \frac{\overline{\partial C_{PM_{2.5},c,r,l=0,t}^{PM_{2.5}}|_t}}{\overline{\partial E_{NO_x,c,r,l,t}|_t}} \overline{E_{NO_x,c,r,l,t}|_t} \quad (15)$$

$$S_{NO_x}^{(2)PM_{2.5}} = \frac{\overline{\partial^2 C_{PM_{2.5},c,r,l=0,t}^{PM_{2.5}}|_t}}{\overline{\partial E_{NO_x,c,r,l,t}^2|_t}} \overline{E_{NO_x,c,r,l,t}^2|_t} \quad (16)$$

Semi-normalised sensitivities reduce the complexity of interpretation by providing sensitivities in the units of the concentration per percent change of emissions. The semi-normalised sensitivities also scale down the impact from cells with low emission rates, which is consistent with the concentration reduction that is realistic **to expect**. Similarly, the time-averaged, semi-normalised cross-sensitivity of $PM_{2.5}$ to both NO_x and monoterpene is denoted as $S_{NO_x,TERP}^{(2)PM_{2.5}}$, with E_{TERP} representing the emission of monoterpenes (Eq. **17**).

$$S_{NO_x,TERP}^{(2)PM_{2.5}} = \frac{\overline{\partial^2 C_{PM_{2.5},c,r,l=0,t}^{PM_{2.5}}|_t}}{\overline{\partial E_{NO_x,c,r,l,t}|_t \partial E_{TERP,c,r,l,t}|_t}} \overline{E_{NO_x,c,r,l,t}|_t \overline{E_{TERP,c,r,l,t}|_t}} \quad (17)$$

It would be good to know the specific cause of the instability from a mathematical viewpoint. Can you derive specifically how the instability grows? This is particularly of interest if the hyd code is truly exact as this would seem to imply some level of inexactness.

- Thank you for this question and perceptive comment. ISORROPIA in reverse mode (with aerosol concentration, RH, and temperature as inputs to the model) is called four times in each execution of the aerosol module in CMAQ in the iterative process of estimating the condensation and evaporation of volatile inorganic gases (HCl , HNO_3 , NH_3) to and from coarse-mode inorganic aerosols (e.g., $ANO3K$, $ANH4K$) to other modes.

Prior research has shown that the aerosol thermodynamic model ISORROPIA in reverse mode leads to unrealistic predictions of changes in aerosol pH and H^+ concentrations (Hennigan et al., 2015). The H^+ concentrations are also extremely sensitive to tiny changes in inputs. This origin of unrealistic sensitivity values is numerically consistent with the underlying equations (based on the ion-balance approach) and, therefore, supports the limits that are based on H^+ to OH^- concentrations.

A second constraint is to limit the changes in H^+ concentrations among the four ISORROPIA runs. After conducting additional testing, we have observed an exponential growth of sensitivity values during one of the first two calls of the reverse mode of ISORROPIA in the coarse mode hybrid equilibration routine in CMAQ. The use of reverse ISORROPIA in CMAQ is similar to a root-finding based approach, and the initial two calls of reverse ISORROPIA lead to unrealistic H^+ concentration changes within a very short timeframe (90 seconds in modeled time). We placed an empirical limit of 1.25 for the changes of H^+ concentration through one single reverse ISORROPIA run for one cell. For instance, if the H^+ concentration changes by 1.25 times from one to the next run, the changes in sensitivity values are ignored. We were still able to achieve the agreement of first- and second-order sensitivity values with respect to the finite-difference or the hybrid-approach with these limits in place.

The hyd code, by mathematical definitions, gives the numerically exact first- and second-order sensitivities of the variable with respect to the perturbed emission value. The numerically exact nature of the hyperdual approach exposes the shortcomings of reverse ISORROPIA.

One important point to note is that despite this shortcoming of ISORROPIA, the performance of CMAQ for inorganic gases and aerosols does not seem to be widely impacted. It is essential to employ the two constraints mentioned in the main manuscript to stop the exponential growth of sensitivities for now. In the future, better constraints and updates to the original model may be able to eliminate the necessity of such conditions.

We have made the description of the two constraints clearer in lines 219 to 223 of the manuscript.

Table 1: The caption needs to state what is being compared. That information can also go directly on the graphs in Fig. 2, so Table 1 is not needed. It would be more effective that way as well.

- Thank you for the helpful direction. We have augmented the caption for Table 1 as indicated below. Unfortunately, since Fig. 2 is a 12-panel plot, we are concerned that the average reader may not be able to see the correlations unless they are included separately as in Table 1. We have also extended the caption of Table 1 and, similarly, of Table 2.

Table 1: The slopes and R^2 values from the linear regression of the first-order sensitivities of ground layer species concentrations of domain-wide perturbations by the hyperdual-step method compared to finite difference sensitivities. The gas-phase species sensitivities with respect to their emissions are line three, where APIN denotes α -pinene and TERP denotes all other monoterpene species. Line four includes sensitivities of aerosol phase products with respect to their precursors where ANO₃ denotes the total aerosol phase nitrate products, ASO₄ denotes the total aerosol sulphate products, and Σ AMT denotes the total aerosol photooxidation products from monoterpene. Line five includes the sensitivities of the total PM_{2.5} concentration with respect to each gas-phase precursor. The visual comparison of the agreement for each relationship is shown in Figure 2.

Table 2: The slopes and R^2 values from the linear regression of the second-order sensitivities of ground layer species concentrations of domain-wide perturbations by the

hyperdual-step method compared to finite difference sensitivities. The gas-phase species sensitivities with respect to their emissions are line three, where APIN denotes α -pinene and TERP denotes all other monoterpene species. Line four includes sensitivities of aerosol phase products with respect to their precursors where ANO₃ denotes the total aerosol phase nitrate products, ASO₄ denotes the total aerosol sulphate products, and Σ AMT denotes the total aerosol photooxidation products from monoterpene. Line five includes the sensitivities of the total PM_{2.5} concentration with respect to each gas-phase precursor. The visual comparison of the agreement for each relationship is shown in Figure 5.

Why not compare the results to another sensitivity analysis method implemented in a CTM, e.g., DDM-3D. This would seem to be much more in line with demonstrating the potential advantages of the method.

- At the time of the development and testing, the DDM-3D for CMAQ version 5.3.2 was not developed yet due to the complexity of updates to a DDM-based approach in complex chemical transport models. The development of HDDM in CMAQ was also compared against a finite-difference-based approach (Zhang et al., 2012). Accordingly, such a comparison is the focus of future work and outside the scope of this manuscript.

Given the description of what was involved, it is not apparent how much of a re-coding savings are involved between the hyd approach and others. Maybe a bit more on the relative effort with more specifics.

- The simplicity of recoding the HYD approach for updates relies on the fact that developers do not need to consider the details about the actual update of the model and construct sensitivity (DDM-3D) or adjoint equations based on the update. Instead, we could change the newly added variable types from “REAL” to “HYPERDUAL”. The sensitivities are calculated line-by-line based on defined hyperdual sensitivities.

To give the reader the sense of this simplicity, we have added the following clause to the last sentence of the first paragraph in the conclusions.

Lines 481 to 482: The development process of CMAQ-hyd is also more straightforward than that of other advanced methods **since all that is needed is to change the type of newly declared variables to hyperdual.**

Did they validate or evaluate the hyperdual module? The two words have rather different meanings.

- Thank you for highlighting this important distinction in language. The reason we use the word “validate” to describe how the hyperdual overloading library was assessed is because these equations are accurate to machine precision. We validated the hyperdual module based on a framework developed by previous work on multicomplex numbers (Pellegrini and Russell, 2016), which facilitates comparison of the hyperdual-based sensitivity to the analytical sensitivity for each of the mathematical functions implemented in the overloading library. We have revised and clarified **lines 194 to 196** of the manuscript.

Lines 194 to 196: Before being applied to CMAQ, the operator overloading library was separately **evaluated against analytical derivatives** using a testing framework developed by Pellegrini and Russell (2016).

One of the more interesting findings of the paper is the computational efficiency found in the hyd method applied to CMAQ vs. other applications. The discuss this a bit, but a bit more analysis would be of interest. For example, for the case of four or eight nodes, say, provide the module-by-module ratio of computational times.

- Thank you for your interest in the detailed computational efficiency of each module. We have developed and now include Figure S3, which demonstrates the relative computational efficiency of each module in the SI of the revised manuscript. We have also revised part of the manuscript to highlight the most interesting findings of profiling, including the relative computational time of the chemistry module (Chem) to the aerosol module (AERO) for CMAQ and CMAQ-hyd.
- In the process of updating this analysis, we identified errant underlying data in the first version of Figure 8. The ‘Other’ category for 1 node, and the MPI_Barrier for 2 nodes have now been corrected, and we have revised Figure 8.
- We have reorganized and added descriptions of the computational time to make the discussion more comprehensive in **lines 456 to 473**.

Lines 456 to 473: With the same computing resources, the total computation time of the CMAQ-hyd is approximately 2.5 (2.44–2.56) times longer. Despite the additional computation burden, CMAQ-hyd remains computationally competitive with the traditional FDM when calculating derivatives. One run of CMAQ-hyd generates the same amount of first- and second-order sensitivity information as at least three runs of regular CMAQ. The relatively low computational cost of CMAQ-hyd, compared to the previous operator overloading approach, may be due to the selective modification of the source code. In contrast to GEOS-CHEM CVM (Constantin and Barrett, 2014), only parts of the model that involve calculating the main species concentration array use hyperdual calculations.

The computational time of scientific modules in CMAQ-hyd generally scales well with increases in computational resources, similar to the original CMAQ. Chem, Aero, and Vdiff are the most computationally expensive modules in both CMAQ and CMAQ-hyd. The relative computational cost of Aero is higher in the CMAQ-hyd than in the regular CMAQ. **The ratio of computational time of Chem to Aero is 1.53 (1.49–1.56) for the CMAQ-hyd runs and 3.98 (3.85–4.19) for the regular CMAQ runs (Fig. 8).** Future work can potentially reduce the computational cost by ignoring sensitivity propagations during the iterative root-finding process in select subroutines, since only the output concentrations from these subroutines are used in the later part of the model. This is also a significant advantage of any operator overloading-based approach (Fike and Alonso, 2011). **The computational time of each module is detailed in Table S3, and the full relative percentage of computational time of each module of eight runs is shown in Figure S3.**

The MPI_Barrier function also scales well with an increasing number of processors. To a certain point, subdividing the domain further reduces the variability of the time required for science processes to be completed across different nodes, resulting in a reduction of the amount of time the program spends waiting for all processes to be synchronized. One important thing to note here is that the scaling of the MPI_barrier is dependent on the number of nodes to number of grid cells.

The I/O process of newly added first- and second-order sensitivity output files increases the computational cost; however, the I/O of species concentration files has a much lower computational cost than other computing modules in CMAQ for this specific scenario. **The I/O processes of CMAQ-hyd and CMAQ take 193 (181–206) seconds and 52 (47–56) seconds, respectively. The I/O process in CMAQ-hyd takes approximately 3.7 times longer on average than that in the regular CMAQ. The overall memory overhead of the CMAQ-hyd is approximately 25 GB for this simulation. A parallel input/output (I/O) approach may be applied to reduce the possibility of potential memory overflow in processor 0 (Wong et al., 2015).**

References

- Fike, J. and Alonso, J.: The Development of Hyper-Dual Numbers for Exact Second-Derivative Calculations, 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, 2011-01-04, 10.2514/6.2011-886,
- Hennigan, C. J., Izumi, J., Sullivan, A. P., Weber, R. J., and Nenes, A.: A critical evaluation of proxy methods used to estimate the acidity of atmospheric particles, *Atmospheric Chemistry and Physics*, 15, 2775-2790, 10.5194/acp-15-2775-2015, 2015.
- Pellegrini, E. and Russell, R. P.: On the Computation and Accuracy of Trajectory State Transition Matrices, *Journal of Guidance, Control, and Dynamics*, 39, 2485-2499, 10.2514/1.G001920, 2016.
- Zhang, W., Capps, S. L., Hu, Y., Nenes, A., Napelenok, S. L., and Russell, A. G.: Development of the high-order decoupled direct method in three dimensions for particulate matter: enabling advanced sensitivity analysis in air quality models, *Geoscientific Model Development*, 5, 355-368, 10.5194/gmd-5-355-2012, 2012.